# Clarifying the role of theory of mind areas during visual perspective taking

Schurz, Matthias; Kronbichler, Martin; Weissengruber, Sebastian; Surtees, Andrew; Samson, Dana; Perner, Josef

*DOI:*
10.1016/j.neuroimage.2015.04.031


# Clarifying the role of theory of mind areas during visual perspective taking: Issues of spontaneity and domain-specificity

Matthias Schurz, Martin Kronbichler, Sebastian Weissengruber, Andrew Surtees, Dana Samson, Josef Perner

PII: S1053-8119(15)00322-5
DOI: doi: 10.1016/j.neuroimage.2015.04.031
Reference: YNIMG 12157

To appear in: *NeuroImage*

Received date: 17 February 2015
Accepted date: 13 April 2015

Please cite this article as: Schurz, Matthias, Kronbichler, Martin, Weissengruber, Sebastian, Surtees, Andrew, Samson, Dana, Perner, Josef, Clarifying the role of theory of mind areas during visual perspective taking: Issues of spontaneity and domain-specificity, *NeuroImage* (2015), doi: 10.1016/j.neuroimage.2015.04.031

# Clarifying the role of theory of mind areas during visual perspective taking: Issues of spontaneity and domain-specificity

Matthias Schurz[1], Martin Kronbichler[1], Sebastian Weissengruber[2], Andrew Surtees[3], Dana Samson[4] & Josef Perner[1].

[1] Centre for Cognitive Neuroscience, University of Salzburg, Austria

[2] Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Switzerland

[3] School of Psychology, University of Birmingham, UK

[4] Psychological Sciences Research Institute, Université catholique de Louvain, Belgium

Correspondence to:

Matthias Schurz.

Postal Address: Centre for Cognitive Neuroscience, University of Salzburg, Hellbrunnerstr. 34, 5020 Salzburg, Austria.

Phone: (+43)/662/8044-5142; Fax:(+43)/662/8044-5126.

E-mail: Matthias.Schurz@sbg.ac.at

**Abstract**

Visual perspective taking is a fundamental feature of the human social brain. Previous research has mainly focused on explicit visual perspective taking and contrasted brain activation for other- versus self-perspective judgments. This produced a conceptual gap to theory of mind studies, where researchers mainly compared activation for taking another's mental perspective to non-mental control conditions. We compared brain activation for visual perspective taking to activation for non-mental control conditions where the avatar was replaced by directional (arrow, lamp) or non-directional (brick-wall) objects. We found domain-specific activation linked to the avatar's visual perspective in right TPJ, ventral mPFC and ventral precuneus. Interestingly, we found that these areas are spontaneously processing information linked to the other's perspective during self-perspective judgments. Based on a review of the visual perspective taking literature, we discuss how these findings can explain some of the inconsistent/negative results found in previous studies comparing other- versus self-perspective judgments.

**Key-words:** Theory of Mind, Mentalizing, Visual Perspective Taking, Spontaneous, Automatic, TPJ, mPFC, Precuneus.

## 1. Introduction

Being able to adopt another's visual perspective is a basic element of human social cognition. It reflects our competence to ascribe a form of mental states - perceptions - to other people. Thus, visual perspective taking (VPT) is sometimes seen as a low level mechanism that is part of our theory of mind (ToM) system (Apperly & Butterfill, 2009). On the neural level, however, the evidence has yet to show a strong link between visual perspective taking and theory of mind.

In an imaging meta-analysis of theory of mind studies, we recently showed that different forms of mental state reasoning (e.g., about beliefs, action goals or personality traits) activate a common set of core-areas comprising of bilateral posterior temporo-parietal junction (TPJp) and dorsal medial prefrontal cortex (mPFC) (Schurz et al., 2014b). The TPJp area was recently distinguished from neighbouring parts of TPJ by a connectivity based parcellation analysis (Mars et al., 2012), and TPJp is located near the Angular Gyrus. When comparing that to a meta-analysis of visual perspective taking studies (Schurz et al., 2013), however, little overlap was found. Most intriguingly, no overlap was found in the right TPJp, which is the most robust neural correlate of theory of mind reasoning. This limited overlap is puzzling in light of conceptual and developmental (e.g., Hamilton et al., 2009) links between visual perspective taking and theory of mind.

In the present study, we want to address a methodological difference between ToM and VPT research which may explains some of the inconsistent findings. VPT imaging research, on the one hand, has so far focused on contrasting self- versus other-perspective judgments (Aichhorn et al., 2006; David et al., 2006, 2008; Dumontheil et al., 2010; Mazzarella et al., 2013; Vogeley et al., 2004; Kockler et al., 2010). ToM imaging research, on the other hand, has contrasted other-mental-perspective judgments versus non-mental control conditions to identify relevant areas (see Schurz et al., 2014b for a review). One argument speaking for non-mental control conditions is that in ToM research activations for self-related judgments overlap with activations for other-related judgments, in particular in mPFC and precuneus (e.g., Mitchell, 2009; Murray et al., 2012). We note, however, that no or only little overlap seems to be present for the TPJ (Saxe et al., 2006b). Similarly, overlaps between self- and other-related judgments were found

across a range of other research domains like pain, disgust, and action (Decety & Grezes, 2006; Keysers & Gazzola, 2009). Therefore, some aspects of mental state reasoning might be cancelled out by a self-versus-other contrast.

The assumption that common processes take place during self- and other-related visual perspective judgments gets additional support through evidence from recent behavioral work. Using a simple (level 1) visual perspective taking task, it was found that participants spontaneously compute the other's perspective when they are asked about what they see themselves in the presence of another person (Samson et al., 2010; Santiesteban et al., 2014; Qureshi et al., 2010). This spontaneous processing linked to the other's perspective was not appropriately picked up in the simple self- versus other-perspective contrast used in most previous research. The present study uses a simple visual perspective taking task modelled after Samson et al. (2010) which allows us to isolate the neural correlates of spontaneous other visual perspective processing by contrasting self-perspective judgments about a scene when an animate object is present in the scene (i.e., a human avatar) compared to when an inanimate object (i.e., an arrow) is present in the scene. We hypothesized that the spontaneous other perspective processing recruits part of the ToM brain network that has usually been washed out in previous studies. Such a finding would provide a way to bridge the gap between ToM and VPT research.

Rather than contrasting other versus self-perspective, we present an arrow as control condition for perspective judgments. Simple arrows are a classic control stimulus in attention research. It was found that simple arrows often produce a similar orientation of attention like the eye gaze (Ristic et al., 2002; Tipples, 2002), and that arrows can produce similar distracting effects on attention as gaze (Nummenmaa and Hietanen, 2009). Recently, it was also hypothesized for spontaneous visual perspective taking that avatars are processed in the same way as arrows simply based on directional features (Santiesteban et al., 2014; for a general review and discussion, see Heyes, 2014). In the present study, the arrow will not only serve as control condition for spontaneous but also for explicit visual perspective taking, where we ask participants a judgments about the object the arrow is pointing at. This makes the arrow an optimal control condition for finding the domain-specific neural mechanism of visual perspective taking. To support the

findings from the comparison of avatar versus arrow, we additionally present a second directional control object (a lamp) and a non-directional control object (brick-wall). These complementary conditions shall elucidate neural processes triggered by the arrow, and if they are qualitatively different from what is found for avatar > arrow.

Another feature of our study is that we completely separate blocks of self and other perspective judgments to ensure that we truly pick up spontaneous visual perspective processing in the brain. In the first run, participants are only asked to tell how many objects were shown on the screen; they are not told about visual perspective and nothing is said about the avatar and the arrow that occasionally occurred in the middle of the screen. This rules out that the other's perspective is made salient by intermixing self- and other-judgments (Samson et al., 2010; Santiesteban et al., 2014). In the second run, participants are asked to explicitly judge the other perspective or tell where the arrow is pointing at.

Put in terms of cognitive theory, our study seeks to determine domain-specificity. By this we mean that the processes for spontaneous and explicit visual perspective taking are distinct from other forms of visuo-spatial reasoning and attention. On the other hand, domain-general would mean that perspective taking relies on visuo-spatial and attentional processes that apply to various task and stimuli, regardless of specific content. In cognitive theory, these assumptions are based on modularity theory (Fodor, 1983, for review see Hirschfeld & Gelman, 1994). In summary, the main aim of our study is to see whether (i) spontaneous visual perspective recruits typical ToM areas and (ii) whether spontaneous and explicit visual perspective taking recruit domain-specific processes.

## 2. Materials and Methods

### 2.1. Participants

The participants for the fMRI study were 22 German-speaking University students (11 male), their mean age was 23.1 years (SD=2.6). All participants were right-handed, and had normal or corrected-to-normal vision and reported no history of neurological or psychiatric disease. Participants gave written

informed consent and received course credit for participation. The study was approved by the ethics committee of the University of Salzburg ("Ethikkommission der Paris Lodron-Universität Salzburg").

2.2. Stimuli and Task

*Main conditions.* The main part of our study was a 2x2x2 factorial experiment with Perspective (self, other), Animacy (avatar, arrow) and Consistency (consistent, inconsistent). We presented a simple level-1 visual perspective task modelled after Samson et al. (2010). In all trials of the experiment, we presented the same basic scenery consisting of a room with discs displayed on the walls. The factor Animacy varied if we additionally presented an arrow pointing either to the left or the right wall or an avatar facing either the left or the right wall (see Figure 1). The factor Consistency varied if the avatar saw / the arrow pointed at the same number of discs that the participant could see herself (consistent) or at a different number of discs than the participant could see (inconsistent; i.e. one or more disc was located on the wall from which the avatar/arrow pointed away). The Perspective factor was varied by two different instructions in different runs. We always asked for self-perspective before other-perspective judgments. For self-perspective, we asked participants to perform a visual object recognition task where they had to indicate how many objects (i.e. discs) were shown in the room (see Figure 1). For the other-perspective task, we asked participants how many objects the avatar in the picture could see. If the picture showed an arrow, we asked how many objects were on the side of the room the arrow was pointing at.

---

Figure 1. Illustration of Task-Design.

---

Stimuli were created with Google Sketchup v3.1 (www.sketchup.com). We presented a male avatar to male participants and a female avatar to female participants. Responses were given by pressing

one out of three buttons on an MRI-compatible response pad, corresponding to the answers "1", "2", or "3".

*Complementary conditions.* In addition, we included two complementary control conditions, which support our main conditions. For self-perspective, we added a lamp and a brick wall condition (both having equal number of trials as our main conditions, i.e. 2 x 42 trials each). For other-perspective, we added the lamp condition, and asked participants on how many discs the light of the lamp is falling if it was turned on (2 x 42 trials).

*Presentation details.* The experiment had an event-related design and the order of stimuli and null-events was determined by a genetic algorithm (Wager & Nichols, 2003) to optimize contrasts between stimulus types (conditions). We presented 42 trials in each condition, resulting in a total of 336 experimental trials plus 168 null-events (1/3 of total trial number) in the self-perspective task and 252 experimental trials plus 126 null-events in the other-perspective task. Each picture was presented for 1500 ms and followed by an interstimulus interval of 1400 ms. Null-events (2900ms) and interstimulus intervals showed the room with a fixation cross in its center (instead of avatar or arrow) and without discs. Participants were asked not to press any button during null-events. Equal numbers of participants were assigned to one of two pseudorandomized stimulus sequences. The order of tasks was fixed, however. Stimulus delivery and response registration were controlled by Presentation (Neurobehavioral Systems Inc., Albany, CA, USA). Before the tasks started, short training sessions were presented (we trained participants separately for self- and other-perspective tasks). All tasks in the scanner took together approximately 45 minutes.

### 2.3. fMRI data acquisition and analysis

Data were acquired on a 3 T MRI scanner (Siemens Magnetom Trio, Siemens Medical Solutions, Erlangen, Germany) using a 32 channel head coil. Functional images were acquired with a T2* weighted echo-planar imaging (EPI) sequence (3x3x3 mm, TR 2250 ms, TE 30 ms, Flip Angle 70°, matrix 64 x 64 mm, FOV 192 mm, Siemens Grappa PAT=2). Within one TR, 36 slices with a slice thickness of 3 mm

were acquired. Data for the self-perspective task was acquired in 2 session with 333 functional images each, the other-perspective task was 2 sessions with 252 images each. In addition, we collected a high-resolution structural image (1 x 1 x 1.2 mm) with a T1-weighted MPRAGE sequence, and a pair of fieldmaps (phase and magnitude images).

Data preprocessing and statistical analysis were performed with SPM 8 (http://www.fil.ion.ucl.ac.uk/spm) running in a MATLAB 7.6 environment (Mathworks Inc., Sherbon MA, USA). Functional images were realigned and unwarped using the parameters from the fieldmap images, and coregistered to the high-resolution structural image. The structural image was segmented into a grey matter, a white matter and a CSF image; then the three images were normalized to corresponding MNI T1 template images. The resulting parameters were used for spatial normalization of the functional images. Functional images were smoothed with a 6 mm FWHM Gaussian Kernel.

Statistical analysis was implemented as a two stage random effects model. In the subject specific first level models, every condition from our experiment was modeled as a separate regressor within a general linear model. These regressors were convolved by a canonical hemodynamic response function. Six regressors coding for head-movements were modeled as covariates of no interest. Functional data in the first levels model were high pass filtered (cut-off 128 sec) and corrected for temporal autocorrelation by an AR(1) model (Friston et al., 2002). At the first level, we calculated linear contrasts between parameter estimates, which reflect signal change for the different conditions versus fixation.

For specific inference, we entered all eight images from each subject into repeated measures ANOVA (flexible factorial design), with the factors Perspective (self, other) and Animacy (avatar, arrow) and Consistency (inconsistent, consistent). Within this model, we computed main effects and interactions. Results were thresholded at a voxelwise $p < .001$ and a cluster-level threshold of $p < .05$ FWE corrected.

To give an overview of how our results relate to those from meta-analyses on VPT (Schurz et al., 2013) and ToM (Schurz et al., 2014b), we show overlaps between maps. No statistics were carried out at this point, as we converted the meta-analysis result maps from Talairach (original) to MNI space by a customized linear transformation which may have produced slight imprecision. To statistically check

overlaps accurately, we additionally carried out ROI analysis based on peak-coordinates from our meta-analyses. Here we could rely on the matrix transformation by Lancaster et al. (2007) to convert meta-analytic peaks from Talairach to MNI.

## 3. Main Results

We analyzed behavioral and imaging data with factorial analyses with the factors Perspective (self versus other), Animacy (avatar versus arrow), and Consistency (consistent versus inconsistent perspective/direction). Because we were interested in domain-specificity and spontaneity of VPT, we first looked at self- and other-perspective tasks separately: For each task, we tested for the main effects of Animacy, the main effect of Consistency, and the interaction between Animacy and Consistency. In a second step, we checked to what extent results for self- and other-perspective judgments were same or different by additionally including the Perspective factor in our analysis. In particular, we looked at effects that included the factor Perspective: Main effect of Perspective, interaction between Perspective and Animacy, interaction between Perspective and Consistency, and triple interaction Perspective by Animacy by Consistency.

### 3.1. Behavioral Results

*Reaction Times.* Figure 2 gives an overview of reaction times on our tasks. First we performed separate ANOVAs for self and other-perspective judgments. For self-perspective judgments the ANOVA showed a main effect of Animacy with longer reaction times for arrow vs. avatar, $F(1,21) = 31.88$, $p < .001$. In addition, we found a main effect of Consistency with longer reaction times for inconsistent trials, $F(1,21) = 9.28$, $p < .01$. The interaction between Animacy and Consistency was not significant ($p > .5$). For other-perspective judgments, we found only the main effect of Consistency to be significant, $F(1,21) = 130.58$, $p < .001$, with longer reaction times for inconsistent trials. The main effect of Animacy and the interaction between Animacy and Consistency were not significant in the other-perspective judgments.

Next, we added the factor Perspective to the ANOVA, making it a 3 factorial analysis. All three main effects were significant. Of main interest for us was the significant Perspective factor, with reaction times slower in other- than self-perspective judgments, $F(1,21) = 15.34$, $p < .001$. Only one of the interactions showed a significant effect, namely Perspective x Consistency, $F(1,21) = 77.26$, $p < .001$. As shown in Figure 2, consistency effects were larger in the other than self-perspective judgments.

*Error Rates.* Error rates were overall low (around 2%) and differences between conditions were very small (1 %). Given these minimal effects, it is unnecessary to carry out statistical tests.

---

Figure 2. Reaction Times and Accuracy.

---

3.2. Imaging Results

3.2.1. Perspective Judgments versus fixation baseline (baseline contrast)

To get an overview of our task-networks, we contrasted brain activation for self and other perspective judgments separately against fixation baseline. To keep it short, we focus on the avatar consistent condition; how results from other conditions relate to this will be addressed in our factorial analysis (next section). Inline Supplementary Figure 1 shows the baseline contrast separately for self perspective (top) and other-perspective (bottom) task. In both perspectives, we found widespread activation in occipital and dorsal occipito-parietal cortices bilaterally. In addition, we found activation on the left side in the inferior parietal lobule, sensorimotor cortex, and parts of the inferior frontal gyrus. We also looked at the overlap of activation maps with results from meta-analyses (shown in white) on VPT (Schurz et al., 2013) and ToM (Schurz et al., 2014b). Large overlaps with the visual perspective taking meta-analysis were present in left occipital, inferior parietal and inferior frontal areas. For the theory of

mind meta-analysis, only marginal overlaps were found bilaterally at the border between temporal and temporo-parietal cortex.

3.2.2. Factorial Analysis

As we have explained before, we first looked at self- and other-perspective tasks separately: We tested for the main effects of animacy, the main effect of consistency, and the interaction between animacy and consistency. The main effect of animacy informs us about domain-specificity, the main effect of consistency – in the self-perspective judgments – informs us about spontaneity of processing another's perspective. Secondly, we checked to what extent results for self- and other-perspective judgments were same or different. Therefore, we carried out a conjunction analysis between the two perspective tasks, and looked at the factor perspective (self versus other) in our design. In particular, we checked for the main effect of perspective, interaction between perspective and animacy, interaction between perspective and consistency, and triple interaction perspective by animacy by consistency. All result-maps concerning the perspective factor were restricted to areas (using inclusive masking) that showed the effect in question in at least one of the perspective-tasks. For example, when looking at the interaction between perspective and animacy, we restricted the analysis to regions that showed an effect for animacy in either self- or other-perspective or in both perspectives. This procedure rules out to find interactions that cannot be clearly linked to an effect in any of the two perspectives, which is of limited interest. All effects were computed by *t*-contrasts, which allowed us to show the direction of main effects and interactions.

3.2.2.1. Results for each perspective separately

*Self-Perspective Task.* As shown in Figure 3, we found activations for the main effect avatar > arrow in right occipito-temporal and temporo-parietal areas, ventral precuneus, and vmPFC/orbitofrontal cortex. Anatomical details are reported in Table 1. For the main effect arrow > avatar, activation was only found in a left occipito-temporal area. The main effect inconsistent > consistent showed activations along

a large stretch of the right occipito-parietal cortex extending to dorsal precuneus, as well as in the cerebellum. The contrast consistent > inconsistent as well as the interaction between animacy and consistency showed no significant results.

*Other-Perspective Task.* The lower section of Figure 3 shows activations for avatar > arrow in the right posterior temporal cortex. Again, see Table 1 for anatomical details. For arrow > avatar, we again found activation in a left occipito-temporal area, and additionally in several visual cortex and dorsal parietal areas. The main effect inconsistent > consistent showed activation in a widespread bilateral occipito-parietal network. The reverse contrast (consistent > inconsistent) and the interaction between animacy and consistency found no significant activations.

---

Figure 3. Result-maps for each perspective separately.

---

*Overlap.* We looked at overlap with meta-analysis maps for theory of mind and visual perspective taking (shown in white in Figure 3). Most interestingly, for the self perspective task, avatar>arrow overlapped with the meta-analytic network for theory of mind in right temporo-parietal cortex, ventral precuneus and mPFC. For the other-perspective task, a marginal overlap was found for avatar>arrow and the theory of mind map in the right posterior temporal cortex. No overlap was found with the meta-analytic map for visual perspective taking.

---

Table 1. Coordinates-results for each perspective separately.

---

3.2.2.2. Conjunction and differences between self- and other-perspectives

*Conjunction analysis.* Figure 4 shows results of a conjunction analysis between self- and other-perspective tasks. We found common activation for avatar > arrow in right posterior middle temporal cortex. For anatomical details see Table 2. For the main effect arrow > avatar, we found common activation in a left occipito-temporal area. For the main effect inconsistent > consistent, we found common activation in right superior and inferior parietal lobes and right middle occipital gyrus.

*Differences between Self- and Other-Perspective Tasks.* We tested all effects in our factorial analysis that involved the perspective factor. First, we tested for the main effect of Perspective. This is an important check for our experiment, since we blocked task-instructions and always asked for self-perspective (i.e. "how many objects are there?") before other-perspective judgments. No brain area showed a significant main effect for Perspective, assuring that our separation of perspectives did not lead to global between-task effects in brain activity.

---

Figure 4. Result-maps for Conj. and Diff. between Self and Other Perspective.

---

Next, we tested for all interactions that included the Perspective factor: Perspective by Animacy, Perspective by Consistency, and the triple interaction Perspective by Animacy by Consistency. Only three interaction effects were significant. First, stronger activation in self-perspective > other-perspective task was found for the main effect avatar > arrow in vmPFC/orbitofrontal cortex and ventral precuneus. Areas are shown in Figure 4. Second, stronger activation for other > self was found for arrow > avatar in right superior parietal lobule and left fusiform gyrus. Third, stronger activation for other > self was found for inconsistent > consistent in bilateral fusiform gyri close to the cerebellum.

*Overlap.* In Figure 4 we also show the overlaps of findings with meta-analysis maps (in white) for ToM and VPT. The most notable overlap with respect to common activations for self- and other-perspective was the conjoint main effect of animacy (avatar > arrow) in the right posterior MTG. This

finding marginally overlapped with the meta-analysis map for ToM. With respect to differences between self- and other-perspective tasks, we observed that findings in vmPFC /orbitofrontal cortex and ventral precuneus overlapped with the ToM meta-analysis map. These areas showed an interaction between animacy and perspective in our analysis, that is, stronger activation in self-perspective > other-perspective for the main effect avatar > arrow.

---

Table 2. Coordinate-results for Conj. and Diff. between Self and Other Persp.

---

3.2.3. ROI Analysis

Our ROI analysis looked at activation in classical theory of mind areas according to literature reviews. Broadly, these are bilateral temporo-parietal cortex and midline areas in mPFC and precuneus (e.g. Amodio & Frith, 2006; Frith & Frith, 2006; Mar, 2011; Mitchell, 2009; Schurz et al., 2014b). Exact coordinates for ROIs were based on peaks from our voxel-wise meta-analyses on ToM and VPT (Schurz et al., 2013; 2014b).

For ROI analyses we carried out the same procedure as for our whole-brain analysis. We performed repeated measures ANOVA with the factors Perspective (self versus other), Animacy (avatar versus arrow) and Consistency (inconsistent versus consistent). Again, we looked at effects for self- and other-perspectives separately and tested for the all main effects and interactions that included the perspective factor. For the sake of brevity, we will only report when effects were found in a ROI, and not mention the effects that were non-significant.

3.2.3.1. Temporo-parietal ROIs

We selected ROIs in bilateral posterior TPJ (TPJp; left: -53, -50, 20; right: 56, -56, 18) based on the ToM meta-analysis (Schurz et al., 2014b) and a ROI in left anterior IPL (IPLa; -37, -44, 47) based on the VPT meta-analysis (Schurz et al., 2013). ROI activation patterns are shown in Figure 5. For the right TPJp, our repeated measures ANOVA found an interaction of Perspective by Animacy, $F(1,21) = 4.8$, $p < .05$. When looking at the perspective judgments separately, we found a marginally significant main effect of Animacy for self, $F(1,21) = 4.3$, $p = .05$, but not for the other perspective. Post-hoc *t*-tests showed that the effect was due to stronger activations in avatar compared to arrow conditions (for details, see Figure 5). For the aIPL, we found effects only when looking at the tasks separately. For the other-perspective task, we found a main effect of Animacy $F(1,21) = 11.24$, $p < .01$, and a main effect of Consistency, $F(1,21) = 9.6$, $p < .01$. Post-hoc *t*-tests showed that effects were driven by stronger activations for arrow>avatar and inconsistent>consistent (see Figure 5).

---

Figure 5. Results ROI Analysis.

---

3.2.3.2. Midline ROIs

In ventral mPFC and ventral precuneus, we found an interaction between Perspective and Animacy, $F(1,21) = 11.9$, $p < .01$ and $F(1,21) = 4.4$, $p < .05$, respectively. When looking at the perspective judgments separately, we found a main effect of Animacy for self-perspective judgments in ventral mPFC $F(1,21) = 14.7$, $p < .001$ and precuneus $F(1,21) = 17.3$, $p < .001$, respectively. Post-hoc *t*-test results in Figure 5 show that effects are driven by stronger activations in avatar > arrow conditions. Completely different patterns were found for dorsal mPFC and precuneus. For dorsal mPFC we found a main effect of perspective $F(1,21) = 7.44$, $p < .01$. Post-hoc *t*-tests showed that activations were stronger in self- compared to other-perspective judgments. For the dorsal precuneus, we found similar main effects when looking at tasks separately. For both self- and other-perspective, we found a main effect of Consistency,

$F(1,21) = 6.6.$, $p < .05$, and $F(1,21) = 14.7$, $p < .001$, respectively. Post-hoc *t*-tests showed stronger activations in inconsistent compared to consistent conditions (see Figure 5).

**4. Complementary Results**

Complementary to our main results, we report data from additional control conditions: a lamp and a brick-wall (i.e., a non-fronted object). Our complementary results are limited to self-judgments, as the issue of domain-specificity is most critical for spontaneous perspective processing in this task. First we look at the lamp condition. Reaction times for the lamp were 794 ms (SD=102) for consistent and 802 ms (SD=108) for inconsistent trials. An ANOVA with the factors Stimulus-Type (arrow, lamp) and Consistency (inconsistent, consistent) found a main effect of Stimulus-Type, i.e. longer reaction times for arrow than for lamp, $F(1,21) = 9.44$, $p < .01$. The main effect of Consistency and the interaction between Stimulus-Type and Consistency were not significant. Second, we compared reaction times for brick-wall (mean=819 ms, SD=101 ms) to arrow. A paired samples *t*-test comparison (averaging over consistent and inconsistent arrows) showed no difference between reaction times, $t < .4$, $p > .7$.

On the neural level, we again compared our complementary control conditions to the arrow. Wherever necessary, we collapsed activation over consistent and inconsistent trials. First, we compared activation for arrow > brick-wall, to show us brain activation for domain-general attention cueing. If activations are similar to what we found for avatar > arrow in our main analysis, this would show that we were only looking at a difference in domain-general attention cueing effects. Figure 6 shows that results for arrow > brick-wall were focused on left inferior parietal lobule and left occipito-temporal cortex, but not present in theory of mind areas identified for avatar > arrow. Details are given in Table 3. The reverse contrast brick-wall > arrow found one cluster of activation in early visual cortex, which may be due to the visually more complex texture on the brick wall.

---

Figure 6. Complementary Results: Arrow versus other control stimuli.

---

Second, we compared activations between arrow and lamp conditions. In particular, we wanted to see if lamp > arrow showed any specific activations. Finding "lamp-specific" activations would challenge the interpretation of "human-specific" activations for the avatar. As shown in Figure 6, we found only one area with stronger activation for lamp > arrow in the right inferior occipital gyrus, similar to what we found for avatar > arrow. The reverse contrast arrow > lamp showed activation in bilateral parietal (left inferior/superior parietal, right superior parietal/middle occipital), left precentral and left occipito-temporal areas.

---

Table 3. Complementary Results: Arrow versus other control stimuli.

---

**5. Discussion**

We tested for spontaneous and domain-specific neural processes in visual perspective taking. Little is known about these topics. Most importantly, no fMRI study has used a directional feature as control condition for perspective taking. We found that three central ToM areas, the right TPJp, vmPFC and ventral precuneus show a coherent result pattern: (1) stronger activation for avatar > arrow, (2) stronger activation for inconsistent > consistent perspectives for avatar, (3) no activation for arrow > non-directional control, (4) no difference in activation between lamp versus arrow control. Taken together, this pattern provides novel insights on neural systems implicated in spontaneous perspective taking. In the following sections, we discuss these findings and their implications in more detail.

5.1. Brain areas engaged in spontaneous perspective taking?

A central feature of our study was that we compared brain activation during self-perspective judgments in the presence of an avatar compared to an arrow. Studies showed that arrows direct or cue attention similar as social stimuli like eye gaze and body posture (Ristic et al., 2002; Tipples, 2002). Therefore, by controlling for the directional features of the avatar, we shed some light on the theory of spontaneous perspective taking.

We found that during self perspective judgments the avatar – but not the arrow - leads to activation in theory of mind areas as identified by a previous meta-analysis (Schurz et al., 2014b). Only based on the location of these areas, however, we cannot say that our participants were properly perspective taking or mentalizing. This conclusion would constitute a case of reverse inference, which has been criticized as invalid if used in a careless way (Hutzler, 2014; Poldrack, 2006). However, our results indicate some functional specificity of these activations found in theory of mind areas. Our main analysis showed that (i) activations were found for an avatar but not for an arrow (main effect of animacy, see Figure 3) and that (ii) activation was also sensitive to the consistency between perspectives – again only for the avatar and not for the arrow (see ROI analysis results in Figure 5).

Further support for functional specificity of activations comes from our complementary analysis. Gaze and body posture of the avatar drive attention just like arrows and other non-human directional cues. An alternative explanation for our avatar>arrow activations in ToM areas could be that they not reflect spontaneous perspective taking due to the agentic features of the avatar but rather different domain-general attentional cueing effects for avatar and arrow. Our complementary contrast arrow > brick-wall is informative about attentional processes due to directional features. The contrast only found activations in left parietal and occipito-temporal areas. Ruling out that activations only reflect different levels of domain-general attention cueing, we found no activation for arrow > brick-wall in ToM areas where we found avatar > arrow.

The contrast lamp > arrow further examined our interpretation. If the contrast lamp > arrow shows activations in the same areas as avatar > arrow, this challenges a “human-specific” interpretation. Our

lamp had a more comparable size and orientation to the avatar, and it is a visually more complex object than an arrow. Furthermore, the lamp's cone of light has spatial similarities to the human's field of view (although we did not show light falling from the lamp in our stimuli). Results showed no activations for lamp > arrow in ToM areas identified by the contrast avatar > arrow, which again corroborates the specificity of activations for avatar > arrow. The only exception to this was that we found activation in the right lateral occipital complex both for lamp > arrow and avatar > arrow. This could be explained by the visual nature of our stimuli: Lamp and avatar are visual objects; arrow is a visual symbol. The lateral occipital complex is specialized for recognition of visual objects (Grill-Spector et al., 2001; Kourtzi & Kanwisher, 2001). Right-sided dominance of overactivation for avatar and lamp could be linked to the fact that visual word and symbol recognition (which applies to the arrow) is left dominant (e.g. Dehaene and Cohen, 2011).

Another notable result of our study is in behavioral data. The finding of altercentric intrusion - i.e. prolonged reaction times on self-perspective judgments when the scene shows someone with a different versus same perspective – is the behavioral key finding that speaks for spontaneous perspective taking (McCleery et al., 2011; Qureshi et al., 2010; Samson et al., 2010, Surtees & Apperly, 2012; Surtees et al., 2013). However, Santiesteban et al. (2014) recently showed that a comparable effect on reaction times can be found for an arrow. This led to the hypothesis that the avatar's directional rather than agentive features drive reaction time differences, showing the need to control for such. In the present study, we presented a simple, canonical arrow, which is the typical control stimulus in attention cueing research. Our arrow failed to produce the same reaction time effect as in Santiesteban et al. (2014) who matched their arrow in terms of visual features to the avatar. Although we also found a trend towards longer reaction times for inconsistent versus consistent arrows, the direct comparison was not significant. Therefore, one can argue that we could have produced a quantitatively stronger cueing effect with an arrow similar to that of Santiesteban et al. (2014). However, as our contrast arrow > brick-wall shows, directional features of the arrow did not lead to activations in ToM areas. Thus, even if we increased directionality of our arrow, this was unlikely to alter results of our comparison with the avatar. Moreover,

we argue that using a canonical arrow makes our imaging results straightforward to interpret. In contrast, interpreting brain activation for an arrow that shares visual features with an avatar (e.g., standing upright and having a directional top) may be less clear. It cannot be ruled out that these features added a rudimentary form of "biological plausibility" or "animacy" to the arrow. This question is beyond the scope of our study and we refrain from taking a strong position here. Future studies are needed to clarify what features exactly are driving the strength of the self-consistency effect for arrows.

5.2. Self versus Other – Relations to previous findings

Our findings raise a new perspective on several previous findings related to other versus self-perspective taking. First, it is of interest that cortical midline areas – vmPFC and ventral precuneus – were engaged in spontaneous other-perspective taking during self-judgments. Previous evidence (e.g. Vogeley et al., 2004) showed activation in midline areas for self-perspective judgments, which was taken to reflect the act of putting oneself in the context of one's environment. This was termed a "state of self" and linked to the brain's "default mode" during passive rest (Vogeley et al., 2003; 2004). In the present study, midline activation was only found for the avatar and not for the arrow. This suggests that activation is not purely based on self-perspective (e.g. "*what do I see?*"), but also depends on presence of another agent. Second, activation was higher when self- and other-perspectives were inconsistent. This further strengthens the idea that midline activation in VPT is not merely self-related, but reflects processing of a self-other perspective difference (e.g., Denny et al., 2012).

Activation in cortical midline regions for self-perspective judgments with the avatar is also interesting for a second reason. As we showed in our direct comparison, this activation was stronger for self compared to other-perspective judgments (in particular for the vmPFC and the ventral precuneus, see Figure 4). This is in line with previous VPT findings, as shown by our review in Figure 7. Among the few studies that did find activation in vmPFC and ventral precuneus, all were using the self>other contrast. It should not be overlooked, however, that our ROI analyses (Figure 5) found activation for the avatar in the other-perspective task. In both right TPJp and ventral precuneus, we found stronger activation for

inconsistent>consistent trials with the avatar (note that this effect was not significant for the arrow). Moreover, we found tendentially stronger activation for avatar inconsistent > arrow inconsistent in other-judgments in ventral precuneus. One possible explanation for reduced activation during other-judgments could be that when participants were explicitly asked to only process a specific feature of the avatar's perspective (how many discs are in his field of view?) they relied on a more simple strategy - like tracing the line of sight – rather than imagining his visual perspective.

As we summarize in Figure 7, previous imaging studies rarely found activation in the right TPJ, vmPFC, and not much more for ventral precuneus. The fact that we found these areas can be explained by difference in task design. When looking at the study descriptions in Figure 7, we can see that nearly all previous studies used the standard contrast other > self-judgment where an avatar with an inconsistent perspective was present in both conditions. In the light of our findings, we suggest that in previous studies spontaneous processing of the difference between self- and other-perspectives probably took place both in the experimental task (other-judgment) and even more so in the control task (self-judgment). Consequently, important activation was blurred or even cancelled out in self versus other comparisons. This seems particularly likely for the rTPJ: Here, our review shows no activation in all studies with the standard contrast other > self-judgment. Only 2 out of 9 studies did find activation in rTPJ, and these were the only ones presenting no avatar (Dumontheil et al., 2010) or an avatar with a consistent perspective (Ramsey et al., 2013) in the control condition. Thus, here control conditions did not promote spontaneous perspective processing.

Our evidence for spontaneous neural processing of visual perspective differences also offers an interesting parallel to recent findings on processing of belief. A number of studies contrasted false belief with true belief reasoning (Aichhorn et al., 2009; Dohnel et al., 2012; Schuwerk et al., 2014a,b; Sommer et al., 2007). Counter to the expectation that rTPJ is processing beliefs per se (e.g., Saxe & Kanwisher, 2003; Saxe & Powell, 2006), studies found stronger activation in rTPJ for false compared to true belief scenarios in verbal and cartoon stories (Aichhorn et al., 2009; Sommer et al., 2007). This was taken to

show that the rTPJ is preferentially activated for mental states that create a perspective difference (Sommer et al., 2007; Perner et al., 2006).

---

Figure 7. VPT Literature Review.

---

More recent studies found equal activation levels in rTPJ for true and false belief when analyzing activation not for mere story comprehension but for answering a question about belief (see Dohnel et al., 2012; Aichhorn et al., 2009). Taken together, findings on belief reasoning concur with our observation that when no explicit prompt for mental or visual perspective taking is given, rTPJ engages only if required by a mismatch between the other's mental state and reality (as we see it). In contrast to this spontaneous but possibly conscious belief processing, recent evidence suggests that rTPJ may not be engaged in implicit belief processing in the absence of awareness (Schneider et al., 2014a; for a theoretical discussion, see Schneider et al., 2014b).

5.3. Other findings

The left IPL is a prominent finding in previous VPT research (see Figure 7) which makes it worth discussing. As shown in Figure 3, we found inconsistent > consistent and arrow > avatar in a left IPL area during other-perspective judgments. A similar pattern was also found during self-perspective judgments. In our complementary analysis, shown in Figure 6, we further found stronger activation in IPL areas for arrow > brick-wall and arrow > lamp. Taken together, these findings implicate that the left IPL is engaged in a domain-general process, irrespective of whether an avatar is seeing or an arrow is pointing at something. This parallels findings from belief research. Here it was found that the left IPL is equally active for reasoning about false belief and for reasoning about false signs (e.g. a false direction sign showed the road to the castle in the wrong way; Aichhorn et al., 2009; Perner et al., 2006).

## 6. Conclusion

We found evidence that visual perspective taking engages theory of mind areas. As we show by comparison with an arrow and two complementary control conditions, these activations are specific for human perspective taking. By systematically varying whether self and other can see the same or not, we find that theory of mind areas engage when the scene shows a perspective difference (i.e. a difference in content between two perspectives). On the contrary, we do not find that theory of mind areas are specifically engaged when participants are asked to adopt the other's perspective. As we lay out in a literature review (see Figure 7), our findings reveal a previously largely unrecognized determinant of visual perspective in the brain, which provides a novel conceptual link to theory of mind research.

## Acknowledgements

This work was financially supported by the Doctoral College "Imaging the Mind" (FWF-W1233). We thank Franziska Pfannerstill for help with the literature review and graphical illustrations.

---

Insert Supplementary Figure 1 here

---

## Figure Captions

Figure 1. Task: In the first run (self), participants only responded according to their own view. In the second run (other), participants indicated what the avatar saw or how many discs were on the side the arrow was pointing at. Responses were given by button press (select among options "1", "2" or "3"). Stimuli: Of main interest were avatar and the arrow stimuli. Intermixed with these stimuli, we also presented a lamp and a brick-wall (as complementary control conditions; see text for more details). In

consistent trials, the avatar / arrow was facing the same number of discs as participants could see. In inconsistent trials, numbers were different.

Figure 2. Behavioral data for main conditions (avatar, arrow) in the self and other task. Means and standard errors are shown. * $p < .05$, ** $p < .01$.

Figure 3. Main results for factorial analysis (avatar vs. arrow, consistent vs. inconsistent) within tasks, i.e. separately for self- and other-perspective judgments. Maps were created at voxel-wise $p < .001$ and cluster-extent threshold of $p < .05$ FWE corrected on the whole brain level.

Figure 4. Main results for factorial analysis (avatar vs. arrow, consistent vs. inconsistent) regarding task-related overlaps and differences. Top: Conjunction between self- and other task. Bottom: Differences between self- and other. Maps were created at voxel-wise $p < .001$ and cluster-extent threshold of $p < .05$ FWE corrected on the whole brain level.

Figure 5. Main results (avatar vs. arrow) for Region of Interest that are taken from Theory of Mind and Visual Perspective Taking meta-analyses. Bars represent scaled brain activity estimates versus fixation baseline (mean, standard errors). * $p < .05$, ** $p < .01$.

Figure 6. Complementary results: Activations for arrow versus other control stimuli (lamp, brick-wall). Maps were created at voxel-wise $p < .001$ and cluster-extent threshold of $p < .05$ FWE corrected on the whole brain level.

Figure 7. Review of existing imaging studies on VPT. Numbers displayed on the brain-slices refer to studies in the table below. We looked for each study if an activation peak was reported for the TPJ or mPFC. TPJ is defined as the sum of AG, SMG, STG and MTG. mPFC is defined as BAs 9 and 10.

**References**

Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., Ladurner, G., 2006. Do visual perspective tasks need theory of mind? Neuroimage 30, 1059-1068.

Aichhorn, M., Perner, J., Weiss, B., Kronbichler, M., Staffen, W., Ladurner, G., 2009. Temporo-parietal junction activity in theory-of-mind tasks: falseness, beliefs, or attention. J Cogn Neurosci 21, 1179-1192.

Amodio, D.M., Frith, C.D., 2006. Meeting of minds: the medial frontal cortex and social cognition. Nat Rev Neurosci 7, 268-277.

Apperly, I.A., Butterfill, S.A., 2009. Do humans have two systems to track beliefs and belief-like states? Psychol Rev 116, 953-970.

Corbetta, M., Shulman, G.L., 2002. Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3, 201–215.

David, N., Aumann, C., Santos, N.S., Bewernick, B.H., Eickhoff, S.B., Newen, A., Shah, N.J., Fink, G.R., Vogeley, K., 2008. Differential involvement of the posterior temporal cortex in mentalizing but not perspective taking. Soc Cogn Affect Neurosci 3, 279-289.

David, N., Bewernick, B.H., Cohen, M.X., Newen, A., Lux, S., Fink, G.R., Shah, N.J., Vogeley, K., 2006. Neural representations of self versus other: visual-spatial perspective taking and agency in a virtual ball-tossing game. J Cogn Neurosci 18, 898-910.

Decety, J., Grezes, J., 2006. The power of simulation: imagining one's own and other's behavior. Brain Res 1079, 4-14.

Dehaene, S., Cohen, L., 2011. The unique role of the visual word form area in reading. Trends Cogn Sci 15, 254-262.

Denny, B.T., Kober, H., Wager, T.D., Ochsner, K.N., 2012. A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. J Cogn Neurosci 24, 1742-1752.

Dohnel, K., Schuwerk, T., Meinhardt, J., Sodian, B., Hajak, G., Sommer, M., 2012. Functional activity of the right temporo-parietal junction and of the medial prefrontal cortex associated with true and false belief reasoning. Neuroimage 60, 1652-1661.

Downing, P.E., Jiang, Y., Shuman, M., Kanwisher, N., 2001. A cortical area selective for visual processing of the human body. Science 293, 2470-2473.

Dumontheil, I., Kuster, O., Apperly, I.A., Blakemore, S.J., 2010. Taking perspective into account in a communicative task. Neuroimage 52, 1574-1583.

Fodor, J.A. 1983. Modularity of Mind: An Essay on Faculty Psychology. Cambridge, Mass.: MIT Press.

Friston, K.J., Glaser, D.E., Henson, R.N., Kiebel, S., Phillips, C., Ashburner, J., 2002. Classical and Bayesian inference in neuroimaging: applications. Neuroimage 16, 484-512.

Frith, C.D., Frith, U., 2006. The neural basis of mentalizing. Neuron 50, 531-534.

Grosbras, M.H., Laird, A.R., Paus, T., 2005. Cortical regions involved in eye move- ments, shifts of attention, and gaze perception. Hum Brain Mapp 25, 140– 154.

Hamilton, A.F., Brindley, R., Frith, U., 2009. Visual perspective taking impairment in children with autistic spectrum disorder. Cognition 113, 37-44.

Heyes, C., 2014. Submentalizing: I am not really reading your mind. Perspect Psychol Sci 9, 131-143.

Hirschfeld, L.A., Gelman, S.A. (Eds.) 1994. Mapping the mind: Domain specificity in cognition and culture. Cambridge: Cambridge University Press.

Hutzler, F., 2014. Reverse inference is not a fallacy per se: cognitive processes can be inferred from functional imaging data. Neuroimage 84, 1061-1069.

Kaiser, S., Walther, S., Nennig, E., Kronmuller, K., Mundt, C., Weisbrod, M., Stippich, C., Vogeley, K., 2008. Gender-specific strategy use and neural correlates in a spatial perspective taking task. Neuropsychologia 46, 2524-2531.

Keysers, C., Gazzola, V., 2009. Expanding the mirror: vicarious activity for actions, emotions, and sensations. Curr Opin Neurobiol 19, 666-671.

Kockler, H., Scheef, L., Tepest, R., David, N., Bewernick, B.H., Newen, A., Schild, H.H., May, M.,

Vogeley, K., 2010. Visuospatial perspective taking in a dynamic environment: perceiving moving objects from a first-person-perspective induces a disposition to act. Conscious Cogn 19, 690-701.

Lancaster, J.L., Tordesillas-Gutierrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., Mazziotta, J.C., Fox, P.T., 2007. Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. Hum Brain Mapp 28, 1194-1205.

Mar, R.A., 2011. The neural bases of social cognition and story comprehension. Annu Rev Psychol 62, 103-134.

Mars, R.B., Sallet, J., Schüffelgen, U., Jbabdi, S., Toni, I., Rushworth, M.F., 2012. Connectivity-based subdivisions of the human right "temporoparietal junction area": evidence for different areas participating in different cortical networks. Cereb Cortex 22, 1894-1903.

Mazzarella, E., Ramsey, R., Conson, M., Hamilton, A., 2013. Brain systems for visual perspective taking and action perception. Soc Neurosci 8, 248-267.

McCleery, J.P., Surtees, A.D., Graham, K.A., Richards, J.E., Apperly, I.A., 2011. The neural and cognitive time course of theory of mind. J Neurosci 31, 12849-12854.

Murray, R.J., Schaer, M., Debbane, M., 2012. Degrees of separation: a quantitative neuroimaging meta-analysis investigating self-specificity and shared neural activation between self- and other-reflection. Neurosci Biobehav Rev 36, 1043-1059.

Nummenmaa, L., Hietanen, J.K., 2009. How attentional systems process conflicting cues. The superiority of social over symbolic orienting revisited. J Exp Psychol Hum Percept Perform 35, 1738-1754.

Perner, J., Aichhorn, M., Kronbichler, M., Staffen, W., Ladurner, G., 2006. Thinking of mental and other representations: the roles of left and right temporo-parietal junction. Soc Neurosci 1, 245-258.

Poldrack, R.A., 2006. Can cognitive processes be inferred from neuroimaging data? Trends Cogn Sci 10, 59-63.

Qureshi, A.W., Apperly, I.A., Samson, D., 2010. Executive function is necessary for perspective selection, not Level-1 visual perspective calculation: evidence from a dual-task study of adults. Cognition 117, 230-236.

Ramsey, R., Hansen, P., Apperly, I., Samson, D., 2013. Seeing it my way or your way: frontoparietal brain areas sustain viewpoint-independent perspective selection processes. J Cogn Neurosci 25, 670-684.

Ristic, J., Friesen, C.K., Kingstone, A., 2002. Are eyes special? It depends on how you look at it. Psychonom Bull Rev 9, 507-513.

Samson, D., Apperly, I.A., Braithwaite, J.J., Andrews, B.J., Bodley Scott, S.E., 2010. Seeing it their way: evidence for rapid and involuntary computation of what other people see. J Exp Psychol Hum Percept Perform 36, 1255-1266.

Santiesteban, I., Catmur, C., Hopkins, S.C., Bird, G., Heyes, C., 2014. Avatars and arrows: implicit mentalizing or domain-general processing? J Exp Psychol Hum Percept Perform 40, 929-937.

Saxe, R., Jamal, N., Powell, L., 2006a. My body or yours? The effect of visual perspective on cortical body representations. Cereb Cortex 16, 178-182.

Saxe, R., Kanwisher, N., 2003. People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". Neuroimage 19, 1835-1842.Saxe & Powell

Saxe, R., Moran, J.M., Scholz, J., Gabrieli, J., 2006b. Overlapping and non-overlapping brain regions for theory of mind and self reflection in individual subjects. Soc Cogn Affect Neurosci 1, 229-234.

Saxe, R., Powell, L.J., 2006. It's the thought that counts: specific brain regions for one component of theory of mind. Psychol Sci 17, 692-699.

Schneider, D., Slaughter, V.P., Becker, S.I., Dux, P.E., 2014a. Implicit false-belief processing in the human brain. Neuroimage 101, 268-275.

Schneider, D., Slaughter, V.P., Dux, P.E., 2014b. What do we know about implicit false-belief tracking? Psychon Bull Rev.

Schurz, M., Aichhorn, M., Martin, A., Perner, J., 2013. Common brain areas engaged in false belief reasoning and visual perspective taking: a meta-analysis of functional brain imaging studies. Front Hum Neurosci 7, 712.

Schurz, M., Kronbichler, M., Crone, J., Richlan, F., Klackl, J., Wimmer, H., 2014b. Top-down and

bottom-up influences on the left ventral occipito-temporal cortex during visual word recognition: an analysis of effective connectivity. Hum Brain Mapp 35, 1668-1680.

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., Perner, J., 2014a. Fractionating theory of mind: a meta-analysis of functional brain imaging studies. Neurosci Biobehav Rev 42, 9-34.

Schuwerk, T., Dohnel, K., Sodian, B., Keck, I.R., Rupprecht, R., Sommer, M., 2014a. Functional activity and effective connectivity of the posterior medial prefrontal cortex during processing of incongruent mental states. Hum Brain Mapp 35, 2950-2965.

Schuwerk, T., Schecklmann, M., Langguth, B., Dohnel, K., Sodian, B., Sommer, M., 2014b. Inhibiting the posterior medial prefrontal cortex by rTMS decreases the discrepancy between self and other in Theory of Mind reasoning. Behav Brain Res 274, 312-318.

Sommer, M., Dohnel, K., Sodian, B., Meinhardt, J., Thoermer, C., Hajak, G., 2007. Neural correlates of true and false belief reasoning. Neuroimage 35, 1378-1384.

Surtees, A.D., Apperly, I.A., 2012. Egocentrism and automatic perspective taking in children and adults. Child Dev 83, 452-460.

Surtees, A., Apperly, I., Samson, D., 2013. Similarities and differences in visual and spatial perspective-taking processes. Cognition 129, 426-438.

Tipples, J, 2002. Eye gaze is not unique: automatic orienting in response to uninformative arrows. Psychon Bull Rev 9, 314-318.

Vogeley, K., Fink, G.R., 2003. Neural correlates of the first-person-perspective. Trends Cogn Sci 7, 38-42.

Vogeley, K., May, M., Ritzl, A., Falkai, P., Zilles, K., Fink, G.R., 2004. Neural correlates of first-person perspective as one constituent of human self-consciousness. J Cogn Neurosci 16, 817-827.

Wager, T.D., Nichols, T.E., 2003. Optimization of experimental design in fMRI: a general framework using a genetic algorithm. Neuroimage 18, 293-309.

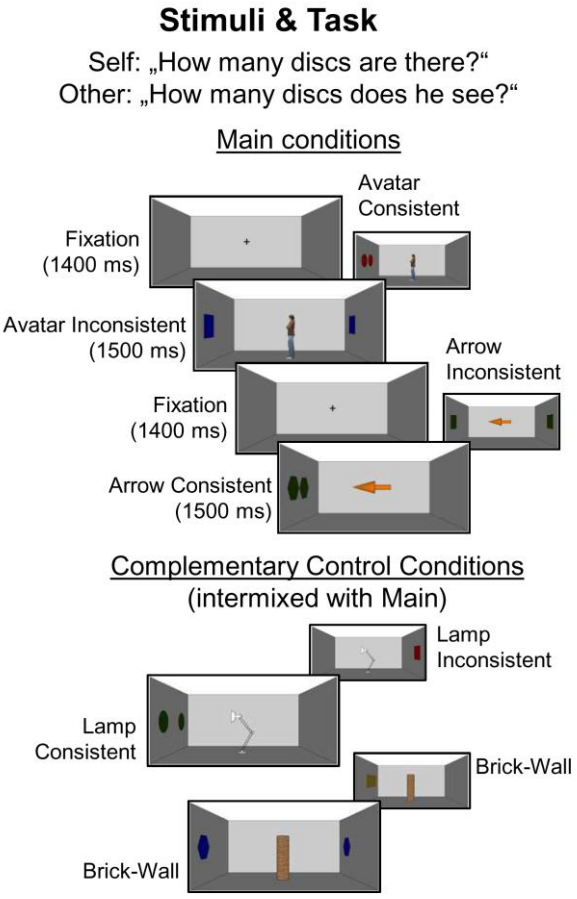



Figure 1

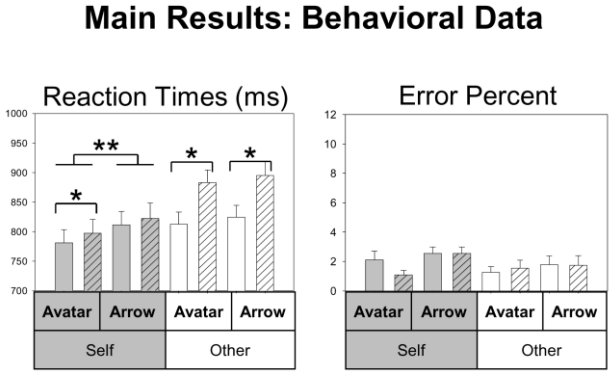



Figure 2

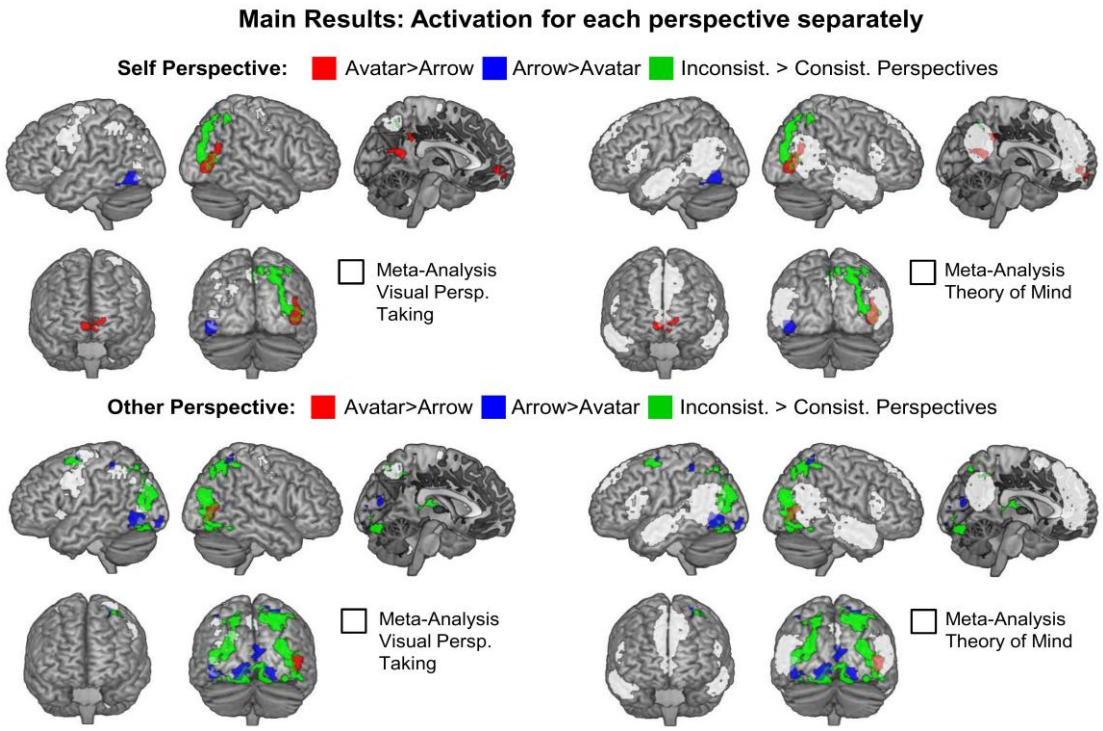



Figure 3

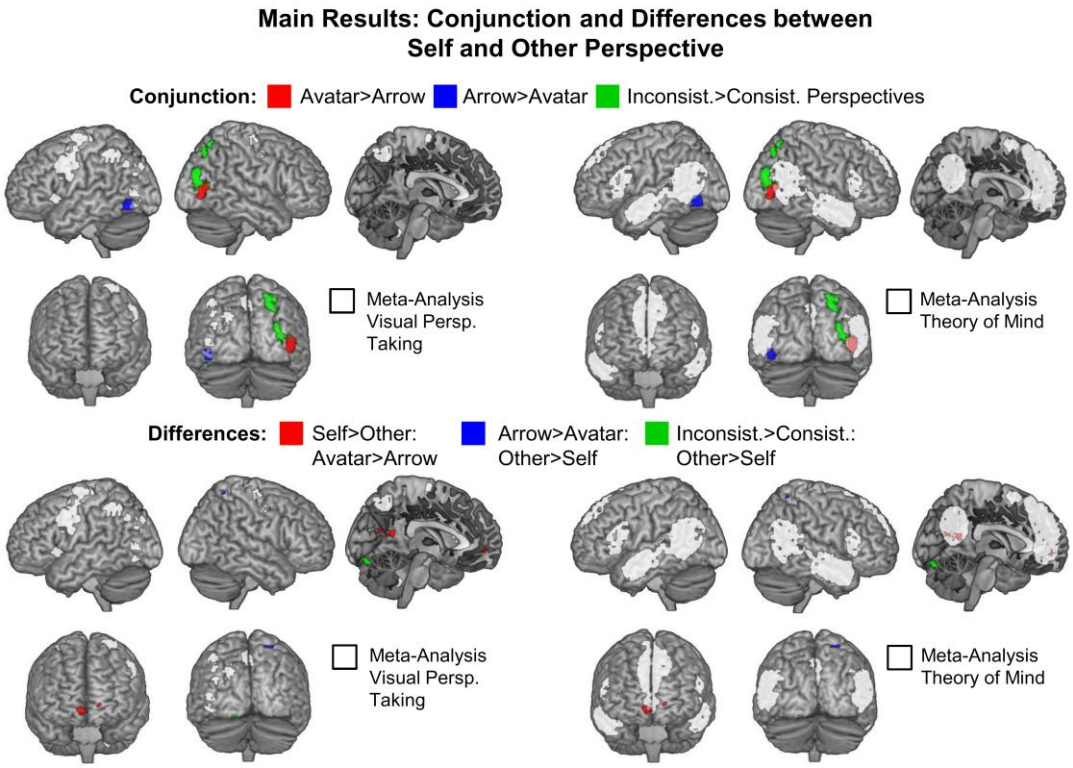



Figure 4

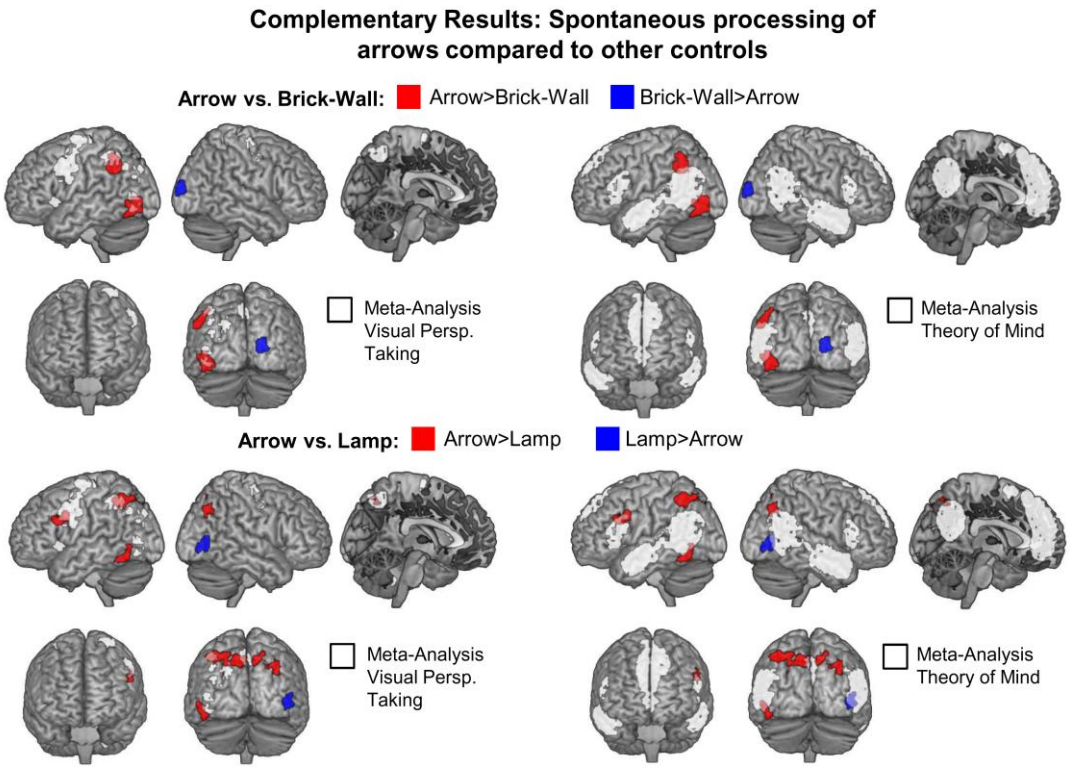



Figure 5

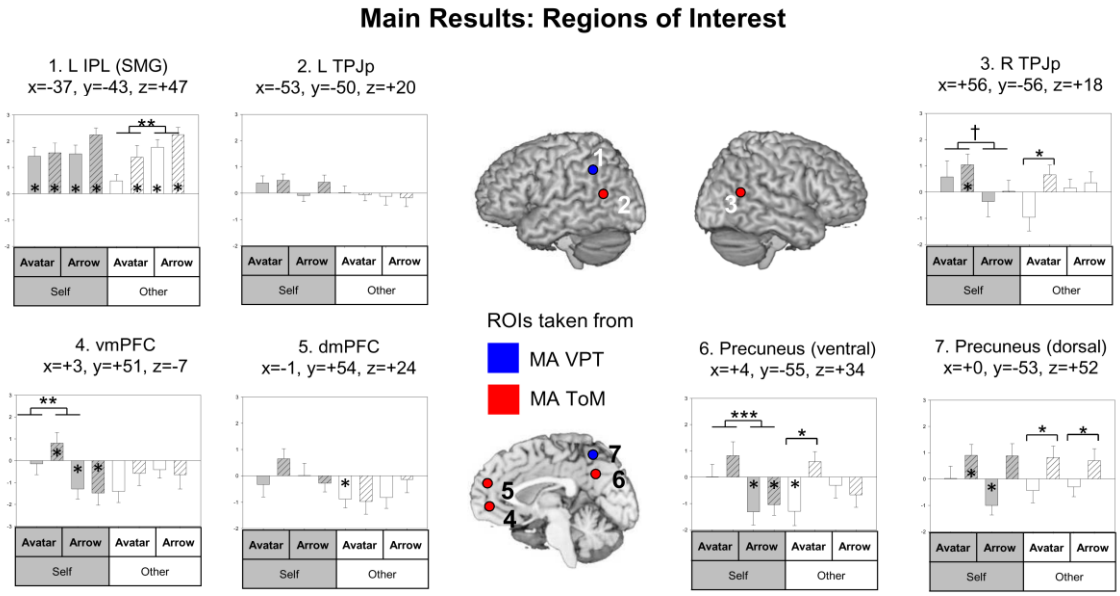
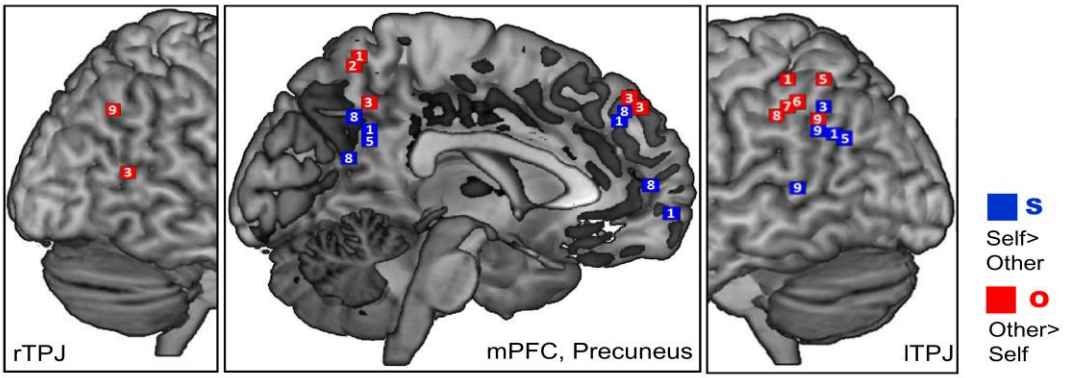



Figure 6

**Review of VPT findings**




| # | L TPJ | R TPJ | M PFC | PRE | Author | Other-judgment: You see… | Self-judgment: You see… |
|---|---|---|---|---|---|---|---|
| 1 | O/S | | S | O | Vogeley 2004 | Avatar and objects. How many does avatar see? | Avatar and objects. How many do you see? |
| 2 | | | | O | Kaiser 2008 | Avatar and objects. How many does avatar see? | Avatar and objects. How many do you see? |
| 3 | S | O | O | O | Dumonth. 2010 | Avatar and objects. Follow avatar's instruction (e.g., 'move the large object up'). Instr. depends on avatar's perspective (e.g. he can't see the largest object). | Only objects. Follow headphone instruction (e.g., 'move the large object up'). Rule: If male (or female) voice, only move particular objects. |
| 4 | | | | | Aichhorn 2006 | Avatar and objects. What's the relative spatial arrangement (e.g. 'Block is in front of the pole') from avatar's viewpoint? | Avatar and objects. What's the relative spatial arrangement (e.g. 'Block is in front of the pole') from your viewpoint? |
| 5 | O/S | | | S | David 2006 | 2 avatars facing you. You play ball-tossing game together. Take perspective of one avatar: Where must he throw the ball (left or right) to pass to other avatar? | 2 avatars facing you. You play ball-tossing game together. Where do you throw the ball (left or right) to pass to one avatar? |
| 6 | O | | | | David 2008 | Avatar facing you; 2 objects in between. Take his perspective: Which object (left or right) is elevated rel. to the other? | Avatar facing you; 2 objects in between. From your perspective: Which object (left or right) is elevated rel. to the other? |
| 7 | O | | | | Kockler 2010 | Avatar and objects. Take his perspective: Which object is to his left-hand side (or right…)? | Avatar and objects. Which object is to your left-hand side (or right…)? |
| 8 | O | | S | S | Mazzar. 2013 | Table with object on it. Avatar stands next to table. Take his perspective: Is object to his left-hand side (or right…)? | Table with object on it. Avatar stands next to table. Is object to your left-hand side (or right…)? |
| 9 | O/S | O | | | Ramsey 2013 | Avatar with inconsistent persp. and objects. How many does avatar see? | Avatar with consistent persp. and objects. How many do you see? |

Figure 7

Table 1. Main Results: Activations for full-factorial design analysis (ANOVA). Main effects of group and interactions between group and task.

| Label | MNI Coordinates | | | *t* | Voxel Extent |
|---|---|---|---|---|---|
| | x | y | z | | |
| **Self** | | | | | |
| **M.E. Avatar>Arrow** | | | | | |
| R mid. occipital | 48 | -76 | 0 | 6.63 | 384 |
| R mid. temporal | 46 | -64 | 20 | 3.98 | - |
| R fusiform / parahipp. | 38 | -36 | -18 | 4.78 | 130 |
| L precuneus / post. cing. | -10 | -50 | 16 | 4.91 | 218 |
| L calcarine | -2 | -66 | 20 | 3.74 | - |
| L mid. cing. / precuneus | -12 | -40 | 36 | 3.94 | 141 |
| L sup. med. front / vmPFC | -16 | 52 | 2 | 4.95 | 304 |
| R orbitofrontal | 6 | 64 | -8 | 4.37 | - |
| **M.E. Arrow>Avatar** | | | | | |
| L inf. occipital | -42 | -70 | -6 | 5.73 | 345 |
| L inf. temporal | -50 | -48 | -18 | 4.32 | - |

**M.E. Inconsistent > Consistent**

| | | | | | |
|---|---|---|---|---|---|
| R sup. parietal | 24 | -74 | 48 | 4.73 | 1275 |
| R mid. occipital | 40 | -80 | 12 | 4.69 | - |
| R mid. temporal | 38 | -62 | 16 | 3.94 | - |
| R inf. parietal | 36 | -46 | 42 | 3.94 | - |
| Cerebellum (vermis) | 6 | -40 | -6 | 4.53 | 159 |
| R parahipp. | 20 | -32 | -14 | 3.88 | - |
| R precuneus | 6 | -54 | 54 | 3.87 | 113 |

Table 1. continued

| Label | MNI Coordinates x | y | z | *t* | Voxel Extent |
|---|---|---|---|---|---|
| **Other** | | | | | |
| **M.E. Avatar>Arrow** | | | | | |
| R mid. temporal | 50 | -70 | 8 | 5.48 | 193 |
| R inf. occipital | 48 | -78 | -4 | 5.04 | - |
| **M.E. Arrow>Avatar** | | | | | |
| R lingual | 16 | -90 | -6 | 5.70 | 191 |
| R sup. parietal | 24 | -54 | 62 | 4.10 | 118 |
| L inf. occipital | -16 | -94 | -10 | 5.70 | 643 |
| L calcarine | 0 | -76 | 16 | 4.13 | - |
| L inf. occipital | -42 | -66 | -4 | 6.23 | 376 |
| L inf. occipital | -44 | -72 | -10 | 5.11 | - |
| L fusiform | -30 | -56 | -12 | 4.48 | 137 |
| L inf. parietal | -32 | -42 | 42 | 4.36 | 209 |
| L inf. parietal | -38 | -40 | 52 | 4.15 | - |

| | | | | | |
|---|---|---|---|---|---|
| L precentral | -28 | -10 | 44 | 5.42 | 246 |
| L sup. frontal | -24 | -4 | 68 | 3.75 | - |
| **M.E. Inconsistent > Consistent** | | | | | |
| R sup. parietal | 24 | -62 | 50 | 5.35 | 1002 |
| R precuneus | 12 | -72 | 56 | 5.17 | - |
| R mid. occipital | 32 | -60 | 34 | 3.53 | - |
| R inf. parietal | 34 | -46 | 40 | 4.30 | - |
| L cerebellum / lingual | -12 | -82 | -16 | 6.26 | 2786 |
| R lingual | 14 | -80 | -14 | 6.20 | - |
| L mid. occipital | -28 | -70 | 30 | 6.30 | 1441 |
| L mid. occipital | -44 | -74 | 14 | 4.87 | - |
| L inf. parietal | -30 | -46 | 46 | 5.18 | 371 |
| L inf. parietal | -42 | -40 | 40 | 4.04 | - |
| L thalamus | -16 | -24 | 14 | 4.08 | 134 |
| L precentral | -36 | -2 | 60 | 4.63 | 134 |

*All results are reported at a voxel-level treshold of p < .001 uncorrected together with a cluster-extent threshold of p< .05 FWE-corrected.*

Table 2. Main Results: Activations for full-factorial design analysis (ANOVA). Main effects of group and interactions between group and task.

| | MNI Coordinates | | | | |
|---|---|---|---|---|---|
| Label | x | y | z | *t* | Voxel Extent |
| **Self AND Other** | | | | | |
| **M.E. Avatar>Arrow** | | | | | |
| R mid. temporal | 50 | -72 | 8 | 5.44 | 160 |
| R inf. occipital | 48 | -78 | -2 | 5.00 | - |
| **M.E. Arrow>Avatar** | | | | | |
| L inf. occipital | -42 | -66 | -6 | 5.41 | 215 |
| **M.E. Inconsistent > Consistent** | | | | | |
| R sup. parietal | 24 | -72 | 50 | 4.62 | 359 |
| R angular | 32 | -58 | 44 | 3.94 | - |
| R mid. occipital | 30 | -60 | 32 | 3.41 | - |
| R mid. occipital | 40 | -82 | 14 | 4.21 | 257 |
| R mid. temporal | 40 | -66 | 14 | 3.44 | - |

**Self > Other**

**M.E. Avatar > Arrow**

| | | | | | |
|---|---|---|---|---|---|
| R orbitofrontal | 8 | 60 | -8 | 4.62 | 84 |
| L sup. med. front. / vmPFC | -2 | -68 | 18 | 4.00 | 63 |
| L ant. cing. | -8 | 50 | 0 | 3.60 | - |
| L calcarine / precuneus | -16 | 52 | 2 | 4.20 | 76 |

**Other > Self**

**M.E. Arrow > Avatar**

| | | | | | |
|---|---|---|---|---|---|
| R sup. parietal | 22 | -54 | 64 | 4.57 | 44 |
| L fusiform | -30 | -54 | -10 | 4.24 | 30 |

**M.E. Inconsistent > Consistent**

| | | | | | |
|---|---|---|---|---|---|
| R fusiform | 30 | -56 | -8 | 4.50 | 45 |
| L cerebellum | -14 | -80 | -16 | 4.48 | 140 |
| L fusiform | -28 | -72 | -14 | 4.04 | - |

Table 3. Complementary Results: Activations from comparison of Arrow versus other control stimuli: Lamp and Brick Wall.

| Label | MNI Coordinates | | | *t* | Voxel Extent |
|---|---|---|---|---|---|
| | x | y | z | | |
| **Arrow > Lamp** | | | | | |
| L inf. parietal | -36 | -56 | 52 | 4.82 | 865 |
| L precuneus | -8 | -68 | 44 | 3.95 | - |
| L inf. parietal | -34 | -48 | 40 | 4.36 | - |
| L sup. parietal | -28 | -68 | 54 | 4.02 | - |
| R mid. occipital | 32 | -64 | 36 | 4.55 | 416 |
| R precuneus | 12 | -70 | 48 | 4.48 | - |
| L precentral | -28 | -2 | 40 | 4.19 | 304 |
| L precentral | -52 | 10 | 34 | 4.06 | - |
| L inf. occipital | -48 | -60 | -16 | 4.04 | 143 |
| L mid. temporal | -56 | -66 | -2 | 3.59 | - |
| **Lamp > Arrow** | | | | | |
| R inf. occipital | 46 | -78 | -2 | 5.34 | 184 |

| | | | | | |
|---|---|---|---|---|---|
| R mid. occipital | 48 | -72 | 6 | 4.21 | - |
| **Arrow>Brick Wall** | | | | | |
| L inf. occipital | -48 | -76 | -2 | 4.99 | 288 |
| L inf. occipital | -38 | -82 | -10 | 4.18 | - |
| L angular | -58 | -54 | 36 | 4.53 | 313 |
| L inf. parietal | -48 | -56 | 46 | 3.96 | - |
| L inf. parietal | -34 | -58 | 42 | | |
| **Brick Wall > Arrow** | | | | | |
| R sup. occipital | 18 | -100 | 6 | 4.78 | 139 |
| R sup. occipital | 22 | -98 | 14 | 4.49 | - |

*All results are reported at a voxel-level treshold of p < .001 uncorrected together with a cluster-extent threshold of p< .05 FWE-corrected.*

Highlights

- We use a novel control condition with an arrow in a VPT task
- Findings clarify meaning of results from classical other versus self comparisons
- Spontaneous other-perspective processing during self-perspective judgments
- ToM areas process inconsistency between perspectives independent of judgment type